JCTC Journal of Chemical Theory and Computation

pubs.acs.org/JCTC Article

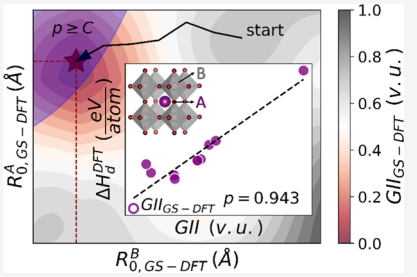
Bond-Valence Parameterization for the Accurate Description of DFT Energetics

Ryan J. Morelock, Zachary J. L. Bare, and Charles B. Musgrave*

Cite This: J. Chem. Theory Comput. 2022, 18, 3257−3267 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: We report a bond-valence method (BVM) parameterization framework that captures density functional theory (DFT)-computed relative stabilities using the BVM global instability index (GII). We benchmarked our framework against a dataset of 188 experimentally observed ABO_3 perovskite oxides, each of which was generated in 11 unique Glazer octahedral tilt systems and optimized using DFT. Our constrained minimization procedure minimizes the GIIs of the 188 perovskite ground state structures predicted by DFT while enforcing a linear correlation between the GIIs and DFT energies of all 2068 competing structures. GIIs based on BVM parameters determined using our framework correctly identified the DFT ground state perovskite structure in 135 of 188 compositions or one of the two lowest energy structures in 152 of 188 compositions. Using the most common approach to parameterize BVM, which minimizes the root-mean-square deviation of the BVM site discrepancy factors, GIIs correctly identified the DFT ground state perovskite structure in only 41 of 188 compositions. Our new parameterization framework is therefore a marked improvement over the existing procedure and an important first step toward BVM-based structure generation protocols that reproduce DFT.



INTRODUCTION

The bond-valence method (BVM) is a localized chemical bonding model based on Pauling's second rule of crystal structures[1] that describes the coordination chemistries of inorganic compounds.[2] For inorganic, ionic crystal structures, Pauling quantified the electrostatic bond strengths, s_{ij}, between cations i and their coordinating anions j as

$$s_{ij} = \frac{z_i}{\text{C. N.}} \tag{1}$$

where z_i are the formal charges (i.e., oxidation states) of cations i, and C.N. are the numbers of anions j constituting the coordination polyhedrons, for example, 4, 6, 8, and so forth. Pauling recognized that many stable crystal structures possess electroneutral local coordination environments such that the oxidation states of anions j (z_j) are the sum of s_{ij} over their coordinating cations i

$$z_j = \sum_i s_{ij} \tag{2}$$

Pauling later showed that z_j is more accurately reproduced by s_{ij} derived from individual cation−anion bond distances rather than the C.N.,[3] and the term bond valence was introduced to differentiate s_{ij} derived in this manner from Pauling bond strengths.[4] In 1978, Zachariasen presented a formalism that relates bond valences s_{ij} to interatomic bond distances, which has been widely adopted by BVM practioners[5]

$$s_{ij} = e^{(R_0 - r_{ij})/B} \tag{3}$$

where R_0 is the ideal bond length of unit valence, which defines the bond length when cation i contributes exactly one valence to anion j. r_{ij} is the bond distance, and B is the softness of the interaction between the cation−anion bonding pair i and j. R_0 and B are the bond-valence parameters, which are fit constants for each cation−anion pair that have units of Å. The bond-valence sum for anions j, BVS_j, adopts the same form as the bond-valence sum for cations i, BVS_i, and is calculated using eq 4

$$\text{BVS}_j = \sum_i s_{ij} \tag{4}$$

The anion site discrepancy factor, d_j, quantifies the deviations between BVS_j and the electroneutral anion coordination environments predicted by Pauling and is calculated using eq 5. For reference, $d_j = 0$ or $d_i = 0$ indicates that the anion or cation is electroneutral when situated in its coordination shell. The cation site discrepancy factor, d_i, is also calculated using eq 5

$$d_j = z_j + \text{BVS}_j \text{ or } d_i = z_i - \text{BVS}_i \tag{5}$$

The global instability index (GII) is defined as the root of the squared discrepancy factors summed over ions k (i.e., both

Received: November 4, 2021
Published: April 20, 2022

https://doi.org/10.1021/acs.jctc.1c01113
J. Chem. Theory Comput. 2022, 18, 3257−3267

anions j and cations i) normalized by the total number of ions present in the crystal structure, N

$$\mathrm{GII} = \sqrt{\frac{\sum_k d_k^2}{N}} \tag{6}$$

For a given composition, the structure with the lowest GII value (in valence units, v.u.) is predicted by BVM to be the most chemically stable, whereas structures with GIIs $\geq$ 0.1 v.u. contain environments that deviate significantly from electroneutrality and are often unstable.[6] The GII thus describes stability as the mean deviation from electroneutrality across all ionic coordination environments, as predicted by the bond-valence parameters R_0 and B.

Calculating the GII is orders of magnitude less computationally demanding than computing stability using quantum mechanics, for example, by density functional theory (DFT) calculations of total energies. Despite their relative simplicity, however, GIIs have been used to accurately describe the DFT energetics of some crystalline solids. For example, Etxebarria et al. related GIIs to the DFT total energies of competing $SrBa_2Ta_2O_9$ structures as a function of distortions along normal modes,[7] while Yamada et al. used GIIs to predict the relative DFT stabilities of eight ABO_3 perovskite compositions as a function of their cubic lattice volumes.[8] The GII's ability to accurately capture DFT energetics depends heavily upon the empirically derived bond-valence parameters R_0 and B, which in turn depend upon both the dataset for which they were parameterized and the parameterization framework used. The tabulated $Co^{3+}-O^{2-}$ ideal bond length parameters R_0 derived for fixed bond softness B = 0.37 Å illustrate the effect of the choice of the dataset; R_0 = 1.637 Å when it was parameterized for a dataset of transition-metal complexes,[9] whereas R_0 = 1.700 Å when it was parameterized for a dataset of crystalline metal oxides.[10] Here, to isolate the effect of the BVM parameterization framework on the GII's ability to describe DFT energetics, BVM parameters were determined for a single, consistent dataset.

ABO_3 perovskite oxides are an ideal dataset to benchmark the GII's ability to describe DFT energetics. Perovskite oxides have been extensively investigated for various applications, including their use as water-splitting redox mediators in renewable solar thermochemical hydrogen (STCH) production.[11−14] Several notable high-throughput DFT investigations of theoretical ABO_3 and alloyed perovskite oxides have been performed to predict the STCH-relevant properties of compositions that have not yet been experimentally synthesized as perovskites.[15−18] Although high-throughput DFT investigations of theoretical perovskites attempt to predict the ground state properties, the DFT ground states of these materials are typically not known before DFT optimization. Consequently, in these investigations the crystal structures and lattice geometries are often assumed and limited to the subsets of known perovskite aristotypes, referred to here as polymorphs. This assumption introduces uncertainty about whether the properties have been calculated for the relevant structure, despite the accuracy of DFT or other higher level methods used to predict the material's properties. For example, Emery and Wolverton computed the properties of perovskite oxide structures in the cubic, rhombohedral, tetragonal, and orthorhombic crystal systems for ∼40% of the compositions considered in their study.[16] However, this study did not explicitly consider the magnitude and phase of perovskite BO_6 octahedral tilting in DFT ground states, which manifests as the Glazer octahedral tilt system,[19] although the DFT-predicted stabilities, band gaps, and optical properties of perovskites are highly sensitive to tilting.[20−22] This deficiency of previous high-throughput investigations could be avoided if BVM parameterization was performed to describe the DFT stabilities of competing perovskite polymorphs, as GIIs calculated from these parameters could correctly identify DFT ground state structures without requiring many demanding DFT optimizations.

The existing BVM parameterization frameworks are ill-suited for deriving R_0 and B BVM parameters that enable GIIs to estimate DFT-predicted relative polymorph stabilities and identify the DFT ground state structures. As discussed by Brown,[23] the most widely used BVM parameterization framework fits R_0 and B vectors that minimize the root-mean-square deviation (rmsd) of squared site discrepancy factors d_k observed in unique coordination polyhedrons n for datasets of well-determined crystal structures,[24] that is, those observed experimentally

$$\min f(\overrightarrow{R_0}, \vec{B}) = \sqrt{\frac{\sum_n \sum_k d_k^2}{n}} \tag{7}$$

$\overrightarrow{R_0}$ and $\vec{B}$ designate vectors of R_0 and B parameters that are fitted for all cation−anion pairs present in the parameterization dataset. The rmsd approach assumes that all structures in the dataset are equally likely to exist experimentally and derives R_0s and Bs that best satisfy site electroneutralities for all coordination polyhedrons present. This assumption is not necessarily valid for the datasets of experimental crystal structures that often contain thermodynamically metastable or unstable phases depending upon their synthesis conditions, as indicated by their calorimetry-measured enthalpies of formation. As shown in the present work, this assumption is inappropriate for the DFT datasets of competing structures, such as ABO_3 perovskite oxides and their polymorphs, which can have vastly different energetics following DFT optimization. For this reason, BVM parameterization based on DFT datasets using the rmsd approach can drastically misorder the stabilities of competing structures using GIIs. This motivated us to develop a new BVM parameterization framework that prioritizes the agreement between GIIs and DFT-computed relative stabilities rather than the R_0 and B vectors' transferabilities across all parameterized structures.

Herein, we report a BVM parameterization framework that minimizes the GIIs of DFT ground state structures (GS-DFT) while enforcing a linear correlation between the GIIs and DFT energies of all competing structures. Our benchmark dataset contains 188 experimental ABO_3 perovskite oxides consisting of 59 unique cation−O^{2-} pairs, which were generated in 11 unique Glazer octahedral tilt systems using the Structure Prediction and Diagnostic Software (SPuDS) program[25] and optimized using GGA + U DFT. DFT relative stabilities are reported as decomposition enthalpies ($\Delta H_d^{\mathrm{DFT}}$) computed relative to the competing phases tabulated in the Materials Project (MP) database.[26] We fixed Bs to be 0.37 Å in our present analysis—as is often the convention when reporting BVM parameters[6]—and parameterized $R_{0,\mathrm{GS\text{-}DFT}}$s using our GS-DFT framework and $R_{0,\mathrm{rmsd}}$s using the existing rmsd framework. We define three parameterization principles used to evaluate the GII's ability to capture DFT relative stabilities, which, to be satisfied, must be attained by most compositions in our dataset

https://doi.org/10.1021/acs.jctc.1c01113
J. Chem. Theory Comput. 2022, 18, 3257−3267

1. GIIs should be physically consistent with the bond-valence model, that is, the GIIs of the DFT-predicted ground state structures ($GII_{GS\text{-}DFT}$) < 0.1 v.u.,
2. $GII_{GS\text{-}DFT}$s should correctly identify ground state structures when GIIs are computed for competing structures, and
3. GIIs should strongly correlate with DFT energies and/or correctly reproduce the DFT energy orderings of competing phases.

Using these criteria, we first evaluated GIIs to describe DFT relative stabilities for the simpler problem of composition-specific $R_{0,GS\text{-}DFT}$s parameterization, where R_0 ideal bond length parameters were derived specifically for each composition. For this case, 376 $R_{0,GS\text{-}DFT}$ parameters (i.e., one $A{-}O^{2-}$ $R_{0,GS\text{-}DFT}$ and one $B{-}O^{2-}$ $R_{0,GS\text{-}DFT}$ for each of the 188 ABO_3 compositions) were derived to describe the energetics of 2068 perovskite oxide structures in our dataset. GIIs based on composition-specific $R_{0,GS\text{-}DFT}$s describe the DFT relative stabilities of competing ABO_3 perovskite polymorphs with exceptional accuracy, validating that $GII_{GS\text{-}DFT}$ minimization with an enforced linear correlation between GIIs and DFT energetics satisfies parameterization principles 1−3. Next, we applied a similar parameterization framework to derive a general set of $R_{0,GS\text{-}DFT}$ parameters. This framework optimized 59 unique values of $R_{0,GS\text{-}DFT}$—for the 59 unique cation$-O^{2-}$ pairs in our dataset—to describe the relative stabilities of all 2068 perovskite oxide structures. While the general BVM $R_{0,GS\text{-}DFT}$ parameters provided by this framework are much more useful, it is significantly more challenging to satisfy parameterization principles 1−3. However, GIIs computed using this general set of $R_{0,GS\text{-}DFT}$ parameters also satisfy all three principles, whereas GIIs based on $R_{0,rmsd}$s only satisfy parameterization principle 1. Finally, we compared the ideal cation$-O^{2-}$ bond distances predicted by the 59 $R_{0,GS\text{-}DFT}$ parameters to those predicted by $R_{0,rmsd}$s to show that they are closer to the actual bond distances observed in DFT-computed ground state structures. This confirms that GS-DFT parameterization derives BVM parameters that enable GIIs to reliably identify the DFT ground state structures of ABO_3 perovskites because the ideal bond lengths predicted by $R_{0,GS\text{-}DFT}$s are more consistent with the bonding environments in DFT ground state structures.

RESULTS

Composition-Specific $R_{0,GS\text{-}DFT}$s Parameterization. General cation−anion bond-valence parameters are conventionally used to compute the GIIs of ionic structures in which cation−anion bonding pairs are present, regardless of those structures' compositions.[27] Thus, the 376 composition-specific $R_{0,GS\text{-}DFT}$ values derived herein and reported in Table S1 are not useful for most practical BVM applications. However, if GIIs computed with the 376 composition-specific $R_{0,GS\text{-}DFT}$s cannot satisfy parameterization principles 1−3 for our dataset, GIIs computed with the 59 general $R_{0,GS\text{-}DFT}$s will certainly not satisfy these criteria. Composition-specific $R_{0,GS\text{-}DFT}$s therefore rapidly establish the upper-performance limit of GIIs to capture DFT energetics using our constrained minimization procedure and confirm the efficacy of this approach. This parameterization was performed for each composition using the following objective function and constraint

$$\min f(\overrightarrow{R_0}, \vec{B}) = GII_{GS\text{-}DFT} \quad (8)$$

such that $p \geq C$

$$p = \frac{\sum_l (GII_l - \overline{GII})(\Delta H_{d,l}^{DFT} - \overline{\Delta H_d}^{DFT})}{\sqrt{\sum_l (GII_l - \overline{GII})^2 (\Delta H_{d,l}^{DFT} - \overline{\Delta H_d}^{DFT})^2}}$$

where l represents the competing structures for each composition or the 11 DFT-optimized Glazer tilts in our dataset. In the present work, all $R_{0,GS\text{-}DFT}$ ideal bond length parameters were parameterized with a fixed bond softness of B = 0.37 Å. Initial $R_{0,GS\text{-}DFT}$ values for the parameterization were derived from Shannon radii[28] and are tabulated in Section II of the Supporting Information.

Composition-specific $R_{0,GS\text{-}DFT}$s minimize $GII_{GS\text{-}DFT}$s subject to the constraint that the Pearson correlation coefficient, p, between GIIs and ΔH_d^{DFT}s for structures l is greater than or equal to the constant, C. The limits on p are that $-1 \leq p \leq 1$, with $p = 1$ indicating a perfect linear correlation between GIIs and ΔH_d^{DFT}s.[29] Because this constraint is only enforced when the total number of competing structures, N^{total}, of a composition is greater than 1, $N^{total} > 1$, and this is true for all 188 compositions in our dataset, the constraint $p \geq C$ was always enforced. We considered several values of C, the relative performances of which are discussed in Section III of the Supporting Information. Minimizing $GII_{GS\text{-}DFT}$s addresses parameterization principle 1, as stable, well-determined crystal structures—or the DFT-predicted ground state structures in DFT datasets—should exhibit $GII_{GS\text{-}DFT}$s < 0.1 v.u. The constraint $p \geq C$ addresses principle 3, as C varies the strength of the positive, linear correlation between GIIs and ΔH_d^{DFT}s, and the upper limit of $C = 1$ forces perfect ordering of ΔH_d^{DFT}s by GIIs for structures of a given composition. Ordering of structures by GIIs is discussed in detail in Section IV of the Supporting Information. Both the objective function and $p \geq C$ constraint address principle 2. The parameterized $R_{0,GS\text{-}DFT}$s should therefore compute the $GII_{GS\text{-}DFT}$s to be less than the GIIs of their competing phases, if these solutions exist.

Figure 1 shows the distribution of $GII_{GS\text{-}DFT}$s resulting from the composition-specific parameterization of $R_{0,GS\text{-}DFT}$s. 183 of 188 compositions (97.3%) exhibit $GII_{GS\text{-}DFT}$s < 0.1 v.u., and GIIs correctly predict the DFT ground state structures for 174 of 188 compositions (92.6%). 186 of 188 compositions (98.9%) exhibit $p \geq 0.7$ for GIIs versus ΔH_d^{DFT}s, indicative of a strong, positive linear correlation. 172 of 188 compositions (91.5%) exhibit $N^{correct} \geq 8$ structures correctly ordered by GIIs relative to ΔH_d^{DFT}s. Composition-specific $R_{0,GS\text{-}DFT}$s using our parameterization framework therefore satisfy parameterization principles 1−3 and show exceptional accuracy in capturing DFT energetics from GIIs, which establishes a high upper-performance limit for general $R_{0,GS\text{-}DFT}$s.

Figure 2 shows the composition-specific GIIs versus ΔH_d^{DFT}s relationships for the eight Co^{3+}-containing ABO_3 compositions in our dataset. All Co^{3+}-containing ABO_3 compositions in our dataset have Co on the perovskite B site and a +3 oxidation state lanthanide cation, Ln^{3+} (Ln^{3+} = Dy^{3+}, Gd^{3+}, Ho^{3+}, Nd^{3+}, Pr^{3+}, Sm^{3+}, Tb^{3+}, and Y^{3+}) on the A site. 28 of the 59 unique cations observed in our dataset are transition-metal cations, including Co^{3+}, whereas 15 of the 59 cations are lanthanides. 102 of 188 ABO_3 compositions have a lanthanide occupying their A site and a transition-metal occupying their B site; therefore, $(Ln^{3+})CoO_3$ compositions are representative of most experimental ABO_3 perovskite oxide compositions in our dataset. For all eight $(Ln^{3+})CoO_3$ compositions ($DyCoO_3$, $GdCoO_3$, $HoCoO_3$, $NdCoO_3$, $PrCoO_3$, $SmCoO_3$, $TbCoO_3$, and $YCoO_3$), GIIs correctly predict the ground state perovskite structure, and

https://doi.org/10.1021/acs.jctc.1c01113
J. Chem. Theory Comput. 2022, 18, 3257−3267

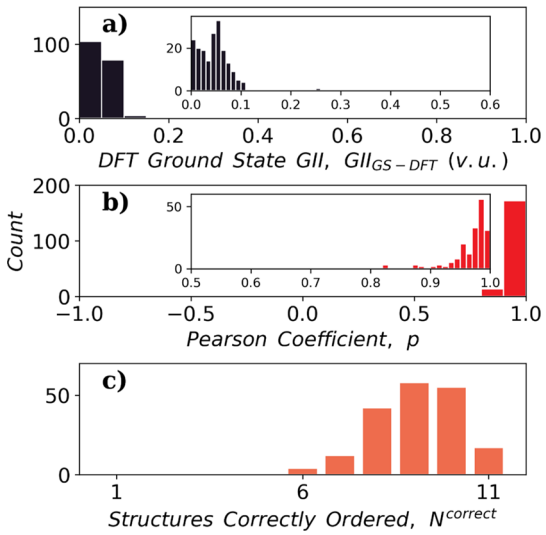


Figure 1. (a) Distribution of DFT ground state global instability indices ($GII_{GS\text{-}DFT}$s) computed using composition-specific $R_{0,GS\text{-}DFT}$s. For 183 of 188 compositions, $GII_{GS\text{-}DFT}$s < 0.1 v.u., which ensures reasonable $R_{0,GS\text{-}DFT}$s. $GII_{GS\text{-}DFT}$s correctly identify the DFT ground states in 174 of 188 compositions. (b) Distribution of composition-specific Pearson coefficients, p. The value of p = 0.232 for $YbAlO_3$ is not shown. (c) Distribution of the number of DFT-optimized Glazer octahedral tilt systems of each composition, with stabilities correctly ordered by GIIs relative to their DFT-predicted decomposition enthalpies, ΔH_d^{DFT}s. Out of 188 compositions, 172 exhibit $N^{correct} \geq 8$ out of $N^{total} = 11$. Parameterization principles 1−3 are therefore satisfied by composition-specific $R_{0,GS\text{-}DFT}$ parameters for most compositions in our dataset.

$GII_{GS\text{-}DFT}$s < 0.1 v.u. ($GII_{GS\text{-}DFT}$s = 0.078, 0.063, 0.078, 0.039, 0.034, 0.068, 0.079, and 0.077 v.u., respectively). As shown in Figure 2, composition-specific $R_{0,GS\text{-}DFT}$s parameterization results in $p \geq 0.7$ for each composition, while the number of correctly ordered structures for each composition is $N^{correct}$ = 7, 7, 6, 7, 8, 9, 8, and 8, respectively. The composition-specific Co^{3+}−O^{2-} $R_{0,GS\text{-}DFT}$ of $GdCoO_3$ ($R_{0,GS\text{-}DFT}$ = 1.699 Å) is noticeably different from Co^{3+}−O^{2-} $R_{0,GS\text{-}DFT}$s of the other parameterized $(Ln^{3+})CoO_3$ compositions, which are grouped near $R_{0,GS\text{-}DFT}$ = 1.768 Å. This suggests that general Co^{3+}−O^{2-} $R_{0,GS\text{-}DFT}$s will not capture DFT relative stabilities from GIIs for the $GdCoO_3$ composition as well as for the remaining $(Ln^{3+})CoO_3$ compositions, an insight that is discussed in Section V of the Supporting Information.

General $R_{0,GS\text{-}DFT}$s and $R_{0,rmsd}$s Parameterization. To derive general $R_{0,GS\text{-}DFT}$ parameters, we modified the composition-specific parameterization objective function and constraint to optimize $R_{0,GS\text{-}DFT}$s for all compositions α containing each $R_{0,GS\text{-}DFT}$ cation−anion pair, again for fixed $\vec{B}$ = 0.37 Å

$$\min f(\overrightarrow{R_0}, \vec{B}) = \sum_{\alpha} w^{\alpha} GII^{\alpha}_{GS\text{-}DFT} \tag{9}$$

such that $\overline{p^{\alpha}} \geq C$

$$\overline{p^{\alpha}} = \frac{\sum_{\alpha} p^{\alpha}}{n^{\alpha}}$$

The constraint on p is enforced for the mean Pearson coefficient computed for all compositions α with the cation−anion pair present, $\overline{p^{\alpha}}$. For the eight Co^{3+} compositions, a single Co^{3+}−O^{2-} $R_{0,GS\text{-}DFT}$ was optimized to minimize the eight $GII_{GS\text{-}DFT}$s and capture the DFT relative stabilities of 88 total GIIs (8 compositions, and 11 competing structures per composition). The median $R_{0,GS\text{-}DFT}$s of composition-specific parameters, $\hat{R}_{0,GS\text{-}DFT}$s, were used as the initial values for general $R_{0,GS\text{-}DFT}$s parameterization. An extensive discussion of $\hat{R}_{0,GS\text{-}DFT}$s, including the distributions from which they were computed, and additional insights gained from composition-specific parameterization prior to general parameterization can be found in Sections V and VI of the Supporting Information. General $R_{0,GS\text{-}DFT}$s parameterization minimizes the weighted sum of $GII_{GS\text{-}DFT}$s ($w^{\alpha}GII^{\alpha}_{GS\text{-}DFT}$) over all compositions α that contain the $R_{0,GS\text{-}DFT}$s' cation−anion pairs. Compositional weighting, that is, variable w^{α}, allows compositions to contribute unequally to the objective function, which can mitigate

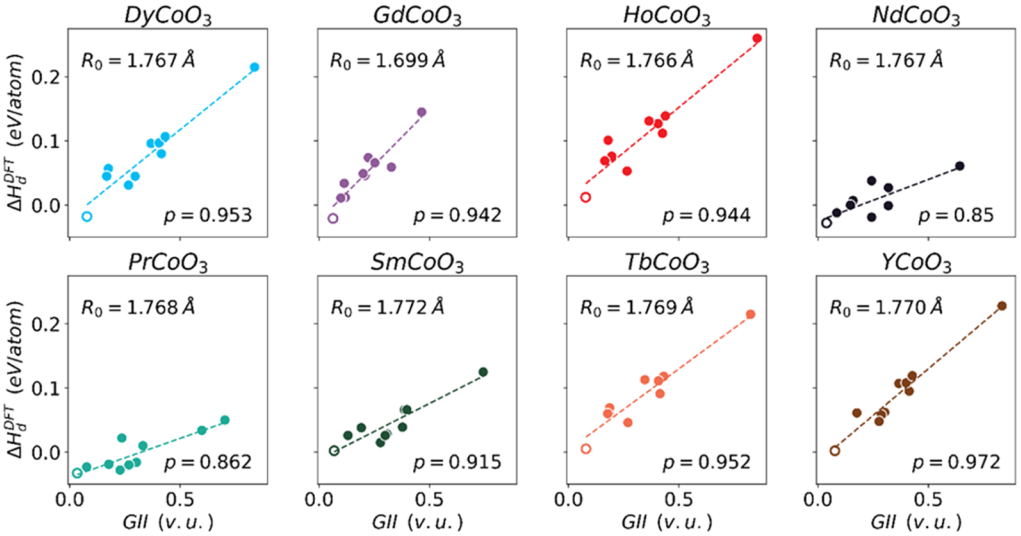


Figure 2. GIIs computed with composition-specific $R_{0,GS\text{-}DFT}$s vs ΔH_d^{DFT}s for the eight lanthanide A site compositions $(Ln^{3+})CoO_3$ in our dataset. Composition-specific Co^{3+}−O^{2-} $R_{0,GS\text{-}DFT}$s for each $(Ln^{3+})CoO_3$, as well as the corresponding Pearson correlation coefficients, p, for the GIIs vs ΔH_d^{DFT}s relationships of the 11 Glazer octahedral tilt systems optimized in DFT, are shown. DFT ground state structures for each composition are represented by unfilled circles, whereas all other structures are represented by filled circles. GIIs computed using composition-specific $R_{0,GS\text{-}DFT}$s correctly predict the DFT ground state structure for all eight $(Ln^{3+})CoO_3$ compositions, all of which exhibit $GII_{GS\text{-}DFT}$s < 0.1 v.u.

https://doi.org/10.1021/acs.jctc.1c01113
J. Chem. Theory Comput. 2022, 18, 3257−3267

overfitting of $R_{0,GS\text{-}DFT}$s to compositions where BVM poorly describes DFT energetics. $R_{0,GS\text{-}DFT}$s parameterization using an example weighting scheme where $GII^{\alpha}_{GS\text{-}DFT}$s were proportionally weighted by $\Delta H^{DFT}_{d,GS}$s is reported in Section VII of the Supporting Information. For our dataset, weighted parameterization does not differ significantly from unweighted parameterization, and therefore no weighting was used. In Section VIII of the Supporting Information, we also show that the results of unweighted, general $R_{0,GS\text{-}DFT}$s parameterization do not change when several Glazer tilt systems are removed from our dataset, validating the robustness of our approach.

Figure 3 shows the distribution of $GII_{GS\text{-}DFT}$s for general $R_{0,GS\text{-}DFT}$ parameters. 130 of 188 compositions (69.1%) exhibit

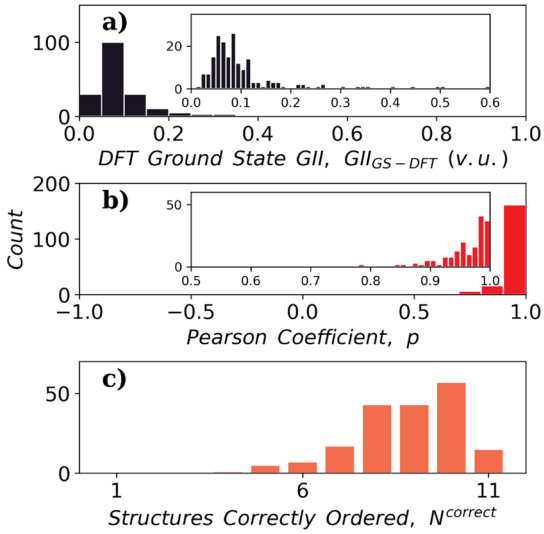


Figure 3. (a) Distribution of DFT ground state global instability indices ($GII_{GS\text{-}DFT}$s) computed using general $R_{0,GS\text{-}DFT}$s. For 130 of 188 compositions, $GII_{GS\text{-}DFT}$s < 0.1 v.u., which ensures reasonable $R_{0,GS\text{-}DFT}$s. $GII_{GS\text{-}DFT}$s correctly identify the DFT ground states in 135 of 188 compositions. (b) Distribution of general Pearson coefficients, p. Values of $p = -0.230$ for $YbAlO_3$, $p = 0.407$ for $YbNiO_3$, and $p = 0.484$ for $AlBiO_3$ are not shown. (c) Distribution of the number of DFT-optimized Glazer octahedral tilt systems of each composition, with stabilities correctly ordered by GIIs relative to their DFT-predicted decomposition enthalpies, ΔH^{DFT}_ds. 158 of 188 compositions exhibit $N^{correct} \geq 8$ out of $N^{total} = 11$. Parameterization principles 1−3 are therefore satisfied by general $R_{0,GS\text{-}DFT}$ parameters for most compositions in our dataset.

$GII_{GS\text{-}DFT}$s < 0.1 v.u., and GIIs correctly predict the DFT ground state structures for 135 of 188 compositions (71.8%). If the DFT ground state structure is one of the two structures with the lowest GIIs, this increases to 152 of 188 compositions (80.9%). 183 of 188 compositions (97.3%) exhibit $p \geq 0.7$ for GIIs versus ΔH^{DFT}_ds. 158 of 188 (84.0%) compositions exhibit $N^{correct} \geq 8$ competing structures correctly ordered by GIIs relative to ΔH^{DFT}_ds. Despite the expected loss in accuracy relative to composition-specific $R_{0,GS\text{-}DFT}$ parameters, general $R_{0,GS\text{-}DFT}$ parameters still satisfy parameterization principles 1−3 for the majority of ABO_3 perovskite oxide compositions. This accuracy is impressive, considering that the mean difference between the ground state ΔH^{DFT}_ds and ΔH^{DFT}_ds of the second lowest energy structures is only 29 meV/atom in our dataset and can be as small as <1 meV/atom, which is within the GGA + U DFT error of ternary oxide stability calculations.[30] Notably, GIIs computed with general $R_{0,GS\text{-}DFT}$ parameters using our approach correctly predict the DFT ground state structure for 75 of 89 compositions (84.3%) with Ln^{3+} on the A site and a 3d transition metal on the B site. This improves to 80 of 89 compositions (89.9%) if the DFT ground state is one of the two structures with the lowest GIIs. This is critically important for high-throughput computational investigations of STCH redox mediators, where complex lanthanide/transition-metal perovskite oxides have proven to be prime candidates.[31,32]

Figure 4 displays the GIIs versus ΔH^{DFT}_ds relationships calculated with general $R_{0,GS\text{-}DFT}$ parameters for the eight $(Ln^{3+})CoO_3$ compositions in our dataset. GIIs correctly predict the ground state crystal structure relative to the competing structures for six of the eight compositions ($DyCoO_3$, $HoCoO_3$, $NdCoO_3$, $PrCoO_3$, $TbCoO_3$, and $YCoO_3$), which exhibit $GII_{GS\text{-}DFT}$s of 0.090, 0.089, 0.056, 0.046, 0.088, and 0.096 v.u. These compositions exhibit values of p of 0.943, 0.932, 0.833, 0.855, 0.942, and 0.956, respectively. $GII_{GS\text{-}DFT}$ is the second lowest GII computed for $SmCoO_3$ (0.080 v.u.) and the fourth lowest computed for $GdCoO_3$ (0.334 v.u.). These compositions exhibit p of 0.907 and 0.722, respectively. The performance of general $Co^{3+}-O^{2-}$ cation−anion $R_{0,GS\text{-}DFT}$ parameters is comparable to that of composition-specific $R_{0,GS\text{-}DFT}$ parameters for seven of the eight $(Ln^{3+})CoO_3$ compositions, which confirms that general $R_{0,GS\text{-}DFT}$ parameters are usually sufficient to describe the DFT relative stabilities of competing ABO_3 perovskite oxide polymorphs. As expected from the results of our composition-specific parameterization, $GdCoO_3$ has the poorest linear correlation between GIIs and the DFT energies. $GdCoO_3$, along with some other compositions considered herein, has competing structures with DFT stabilities that are equal to or within DFT error and, consequently, can be difficult to resolve with the GII. $GdCoO_3$ stabilities are also heavily dependent upon magnetic spin ordering in DFT.[33] Stabilization due to magnetism is considered during the $R_{0,GS\text{-}DFT}$s parameterizations discussed herein, as magnetic ordering can change the ΔH^{DFT}_ds of DFT-optimized structures, but different $R_{0,GS\text{-}DFT}$ parameters for high-spin versus low-spin cations were not explicitly derived in the present work.

We report the 59 general $R_{0,GS\text{-}DFT}$s derived from our dataset in Table S6 of the Supporting Information, along with the 59 general $R_{0,rmsd}$s derived using eq 7. Histograms displaying the results of $R_{0,rmsd}$s parameterization are shown in Figure S6 of the Supporting Information. 122 of 188 compositions (64.9%) have $GII_{GS\text{-}DFT}$s < 0.1 v.u. when computed with $R_{0,rmsd}$s. GIIs correctly identify the ground states in 41 of 188 compositions (21.8%), and 54 of 188 compositions exhibit $N^{correct} \geq 8$ (28.7%). rmsd parameterization using eq 7 therefore satisfies parameterization principle 1 for most ABO_3 perovskite oxide compositions in our dataset but not principles 2 and 3. This behavior is expected, as rmsd parameterization prioritizes R_0 and B vectors' transferabilities between structures over agreement between GIIs and DFT relative stabilities. Minimizing the summed site discrepancy factors produces R_0s and Bs such that the coordination environments of all structures in the dataset (including the ground states) approach electroneutrality, which satisfies principle 1. However, rmsd parameterization does not explicitly consider the GIIs versus ΔH^{DFT}_ds relationship and is therefore unlikely to inherently satisfy parameterization principles 2−3, in contrast to our framework.

Comparison of $R_{0,GS\text{-}DFT}$s to $R_{0,rmsd}$s. Differences in the GII's ability to describe DFT energetics result from the differences between the parameterized values of $R_{0,GS\text{-}DFT}$ and $R_{0,rmsd}$, which in turn affect the ideal cation−anion bond lengths

https://doi.org/10.1021/acs.jctc.1c01113
J. Chem. Theory Comput. 2022, 18, 3257−3267

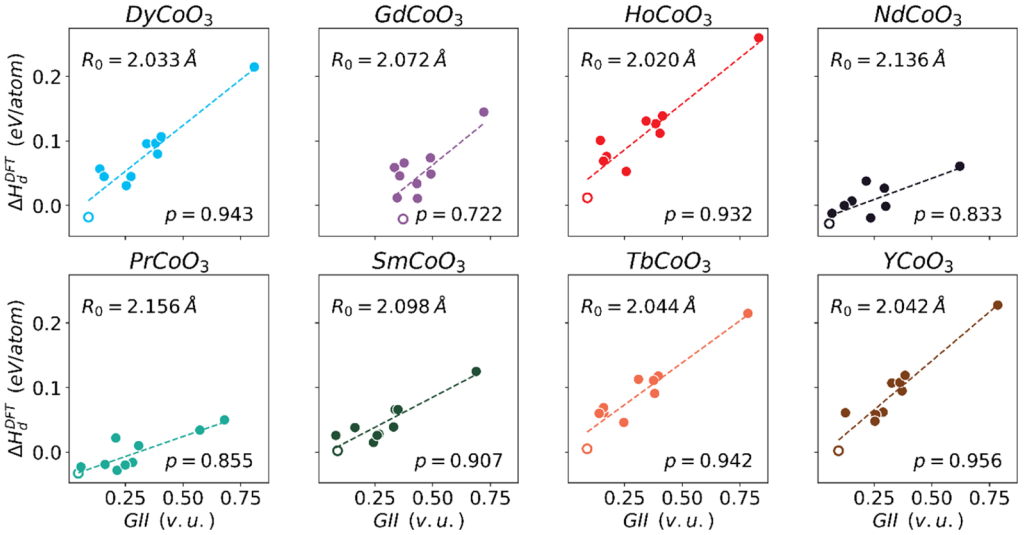


Figure 4. GIIs computed with general $R_{0,\text{GS-DFT}}$s vs ΔH_d^{DFT}s for the eight $(Ln^{3+})CoO_3$ compositions in our dataset. The general $Co^{3+}-O^{2-}$ $R_{0,\text{GS-DFT}}$ parameter optimized using our framework is $R_{0,\text{GS-DFT}}$ = 1.763 Å. General $Ln^{3+}-O^{2-}$ $R_{0,\text{GS-DFT}}$s for each $(Ln^{3+})CoO_3$ and the corresponding coefficients p for the GIIs vs ΔH_d^{DFT}s relationships of the 11 Glazer octahedral tilt systems optimized in DFT are shown. DFT ground state structures are represented by open circles, whereas all other structures are represented by filled circles. GIIs computed using general $R_{0,\text{GS-DFT}}$s correctly predict the DFT ground state structure for six of the eight $(Ln^{3+})CoO_3$ compositions, all six of which exhibit $\text{GII}_{\text{GS-DFT}}$s < 0.1 v.u.

predicted by BVM. These ideal bond lengths can be computed using eq 10

$$r_{ij}^{\text{BVM}} = R_0 - B \ln\left(\frac{z_i}{\text{C. N.}}\right) \tag{10}$$

which is obtained by combining eqs 1 and 3. Here, r_{ij}^{BVM} is the cation−anion bond length predicted by BVM, z_i is the oxidation state of cation i, and C.N. is the number of anions j coordinating cation i. Because $R_{0,\text{GS-DFT}}$s are derived to minimize $\text{GII}_{\text{GS-DFT}}$s, the r_{ij}^{BVM} values of $R_{0,\text{GS-DFT}}$s should better describe bonding environments present in the DFT ground state structures compared to less stable, competing structures. In contrast, the r_{ij}^{BVM} values of $R_{0,\text{rmsd}}$s should not necessarily describe the ground state bonding environments better than the least stable structures, as this parameterization does not differentiate between ground state and metastable structures. To evaluate this assumption, we compared the r_{ij}^{BVM} values predicted by $R_{0,\text{GS-DFT}}$s and $R_{0,\text{rmsd}}$s to the average bond distances, $\bar{r}_{ij}$, observed in the DFT ground state structures versus the least stable structures for all ABO_3 compositions in our dataset. C.N. used to compute r_{ij}^{BVM}s were taken as the number of nearest-neighbor anions within each cation's first coordination sphere. Because the irreducible representations of the 11 Glazer tilts in our dataset range from 5 to 20 atoms, we normalized the ground state and least stable structures to a total of 20 atoms (formula unit $A_4B_4O_{12}$) to ensure that all cation coordination environments in our dataset were counted equally. This resulted in 8 cation sites per structure and a total of 1504 (8 cation sites and 188 ground state structures per composition) coordination environments analyzed for the ground state and least stable structures, respectively.

Figure 5 confirms that $R_{0,\text{GS-DFT}}$s describe the cation−anion bonding environments of the DFT ground states relative to the DFT least stable states better than $R_{0,\text{rmsd}}$s, as indicated by the significant differences in $|r_{ij}^{\text{BVM}} - \bar{r}_{ij}| \leq 0.02$ Å for the ground states versus least stable structures. Furthermore, the differences between the general $R_{0,\text{GS-DFT}}$ bond length parameters and $R_{0,\text{rmsd}}$s are systematic, as shown in Figure S7, where the ideal

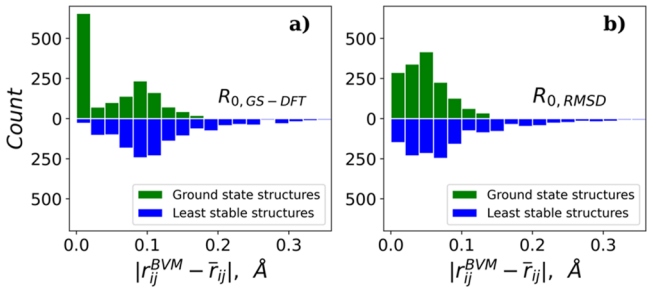


Figure 5. Absolute differences between r_{ij}^{BVM}s and $\bar{r}_{ij}$s for the DFT ground state structures (green) and least stable structures (blue) for r_{ij}^{BVM}s computed from (a) $R_{0,\text{GS-DFT}}$s and (b) $R_{0,\text{rmsd}}$s. r_{ij}^{BVM}s computed from $R_{0,\text{GS-DFT}}$s result in 656 of 1504 ground state cation sites with $|r_{ij}^{\text{BVM}} - \bar{r}_{ij}| \leq 0.02$ Å, indicating that r_{ij}^{BVM}s predicted by $R_{0,\text{GS-DFT}}$s accurately represent the average bond distances, $\bar{r}_{ij}$, present in these coordination environments. This is in sharp contrast to the least stable structure cation sites, for which r_{ij}^{BVM}s computed from $R_{0,\text{GS-DFT}}$s result in 29 of 1504 sites, with $|r_{ij}^{\text{BVM}} - \bar{r}_{ij}| \leq 0.02$ Å. This discrepancy is not nearly as large as r_{ij}^{BVM}s computed from $R_{0,\text{rmsd}}$s, with 288 of 1504 cation sites in the ground state structures and 149 of 1504 cation sites in the least stable structures exhibiting $|r_{ij}^{\text{BVM}} - \bar{r}_{ij}| \leq 0.02$ Å.

bond length $R_{0,\text{GS-DFT}}$s and $R_{0,\text{rmsd}}$s of the two parameterizations are compared based on the cation's group in the periodic table. $R_{0,\text{GS-DFT}}$s are consistently larger than $R_{0,\text{rmsd}}$s for the 28 transition-metal cation−anion pairs, as indicated by their mean signed deviation in Table S7 (MSD = 0.021 Å), and are consistently smaller for the 15 lanthanide cations (MSD = −0.028 Å). $R_{0,\text{GS-DFT}}$s are also smaller than $R_{0,\text{rmsd}}$s for alkali metal and alkaline earth metal cation−anion pairs (MSD = −0.023 Å and MSD = −0.03 Å, respectively), but fewer of these pairs exist in our dataset (only Na^+-O^{2-} and $Mg^{2+}-O^{2-}$, $Ca^{2+}-O^{2-}$, $Sr^{2+}-O^{2-}$, and $Ba^{2+}-O^{2-}$). Because lanthanides, alkali metals, and alkaline earth metals typically occupy the perovskite oxide A site, whereas transition metals typically occupy the B site, differences between parameterized $R_{0,\text{GS-DFT}}$s and $R_{0,\text{rmsd}}$s must also systematically affect the ideal A−O and B−O bond lengths predicted by BVM.

https://doi.org/10.1021/acs.jctc.1c01113
J. Chem. Theory Comput. 2022, 18, 3257−3267

The ideal perovskite A−O and B−O bond distances predicted by BVM, r_{AO}^{BVM} and r_{BO}^{BVM}, were computed using eq 10, for which we obtained C.N. by assuming that the A site cations are 12-fold-coordinated by the X site O^{2-} anions and the B site cations are 6-fold-coordinated, as is the case for the cubic perovskite symmetry. We evaluated the differences between r_{AO}^{BVM}s and r_{BO}^{BVM}s predicted by $R_{0,GS\text{-}DFT}$s and $R_{0,rmsd}$s using the bond-valence tolerance factor, t_{bv}, shown in eq 11. t_{bv} is computed using the same expression as the Goldschmidt tolerance factor[34]—which predicts the ABX_3 compositions that form perovskites—and therefore the perovskite/nonperovskite classifications of the Goldschmidt tolerance factor are expected to apply to t_{bv} as well. However, t_{bv} is computed using r_{AO}^{BVM}s and r_{BO}^{BVM}s, whereas the Goldschmidt tolerance factor is computed using the other estimates of A−O and B−O bond lengths, for example, Shannon ionic radii

$$t_{bv} = r_{AX}^{BVM}/\sqrt{2}\, r_{BX}^{BVM} \quad (11)$$

As reported by Goldschmidt, ABX_3 compositions with $0.8 \leq t_{bv} \leq 1$ typically form the perovskite phase. Compositions approaching $t_{bv} = 1$ are more likely to form the undistorted cubic perovskite with isometric crystal symmetry due to the larger A site cations in these compositions, which occupy more interstitial space between BX_6 octahedra and suppress octahedral tilting or rotation. In contrast, compositions approaching $t_{bv} = 0.8$ have smaller A site cations and are more likely to be stabilized by octahedral tilting, adopting crystal symmetries other than the isometric, that is, orthorhombic, rhombohedral, and so forth.

Figure 6 compares t_{bv}s computed from $R_{0,GS\text{-}DFT}$s to those computed from $R_{0,rmsd}$s for the 188 ABO_3 perovskite compositions in our dataset. Each point represents a different perovskite composition and is colored based on the stabilization due to BO_6 tilting, which we define as the energy difference between that composition's undistorted cubic perovskite phase (Glazer tilt a0a0a0) and its DFT ground state perovskite phase. As discussed in the Methods section, the SPuDS software program used to generate initial structures for DFT optimization produces all 10 remaining Glazer tilts with some degree of BO_6 octahedral tilting. Additionally, all 188 compositions have a DFT-predicted ground state other than the cubic perovskite, such that the stabilization due to BO_6 tilting is negative, that is, some amount of BO_6 tilting is energetically preferred, for all 188 compositions. As shown in Figure 6, systematic shifts in $R_{0,GS\text{-}DFT}$s for the lanthanides and transition metals result in t_{bv}s that predict octahedral tilting to be more energetically favorable (i.e., t_{bv} closer to 0.8) relative to t_{bv}s computed from $R_{0,rmsd}$s. This is quantified by the shift in median t_{bv} values (0.901 for $R_{0,GS\text{-}DFT}$s vs 0.933 for $R_{0,rmsd}$s), which agrees better with the energetic stabilization from BO_6 tilting observed for all compositions in our DFT dataset. We conclude that parameterizing BVM to consider DFT relative stabilities not only improves the GII's predictions of the ground state and crystal structure orderings, but it also yields $R_{0,GS\text{-}DFT}$ ideal bond length parameters that more accurately describe the coordination environments of ground state structures, resulting in strictly structural relationships such as t_{bv} that better agree with DFT energetics.

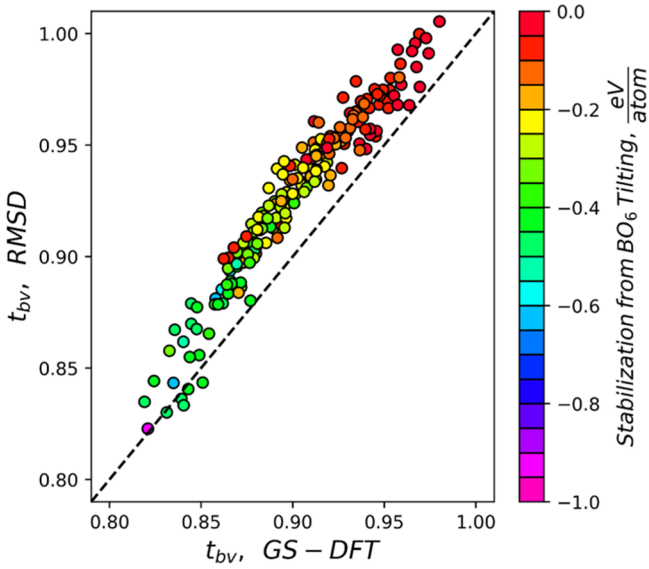


Figure 6. Parity plot comparing t_{bv}s computed from ideal bond lengths predicted by $R_{0,GS\text{-}DFT}$s (x-axis) to those predicted by $R_{0,rmsd}$s (y-axis) for the 188 compositions in our ABO_3 perovskite oxide dataset, all of which exhibit octahedrally distorted DFT ground states that are more stable than the cubic perovskite. Energy differences between the DFT-optimized cubic perovskite and octahedrally distorted DFT ground states (in eV/atom) are indicated by the color bar. Stabilization from BO_6 tilting is generally more preferred (becomes more negative) as t_{bv} approaches 0.8, confirming the observations of Goldschmidt. t_{bv}s computed from $R_{0,GS\text{-}DFT}$s ($0.819 \leq t_{bv} \leq 0.980$) are shifted toward $t_{bv} = 0.8$—where octahedral tilting is more energetically favored—relative to t_{bv}s computed from $R_{0,rmsd}$s ($0.823 \leq t_{bv} \leq 1.005$), in agreement with DFT energetics.

DISCUSSION

The BVM parameterization reported herein optimizes $R_{0,GS\text{-}DFT}$s at a fixed bond softness parameter, B, of 0.37. However, optimizing both $R_{0,GS\text{-}DFT}$ and B values is a relatively straightforward extension that could improve the GII's ability to capture the trends in DFT energies.[35] GIIs could be further modified to include anion contributions beyond the first coordination sphere, otherwise known as softBV parameters,[36] and cation−cation and/or anion−anion interactions.[37] Additionally, high-spin versus low-spin $R_{0,GS\text{-}DFT}$ parameters could be derived for magnetic elements,[38] as these modifications are readily accommodated by the parameterization framework presented herein. Although general $R_{0,GS\text{-}DFT}$ parameters sufficiently describe DFT energies for most experimental ABO_3 perovskite oxides, if more accurate descriptions of competing phase stabilities are desired, we recommend either modifying the BVM framework in these ways and/or using composition-specific BVM parameters, which we have shown to be exceptionally accurate for describing DFT energetics with GIIs. Our parameterization framework's success in describing ABO_3 perovskites is encouraging for other chemical spaces, which we expect can be characterized with similar accuracy.

As demonstrated by the shift in t_{bv} values, the $R_{0,GS\text{-}DFT}$ ideal bond length parameters predict greater energetic stabilization from octahedral tilting than $R_{0,rmsd}$s for the perovskite oxides in our dataset, in agreement with DFT. The structure prediction software, SPuDS, that was used to generate Glazer tilts for the present investigation could potentially use $R_{0,GS\text{-}DFT}$ parameters to predict perovskite structures that more accurately capture BO_6 tilting relative to DFT. This could vastly accelerate high-throughput computational investigations of compelling perovskites beyond the oxides, such as the halide perovskites,[22,39,40] by reducing the computational expense associated with optimizing poor initial structures and modeling multiple

https://doi.org/10.1021/acs.jctc.1c01113
J. Chem. Theory Comput. 2022, 18, 3257−3267

structures that are not the DFT ground state. SPuDS for DFT would all but eliminate uncertainties in property predictions that arise when the correct ground state is not modeled. As demonstrated by Xie et al., SPuDS correctly identifies the ground state Glazer tilt mode from GII in only 9 of 60 perovskite oxide compositions considered[41] and is therefore not currently adept at identifying the correct ground state structures from GIIs. $R_{0,GS\text{-}DFT}$s derived to consider DFT energetics could, in principle, compute GIIs that more accurately generate and order the perovskite Glazer tilts produced by SPuDS. This is a complex problem that warrants its own investigation.

One issue specific to SPuDS is that its optimization algorithm imposes additional constraints, which include fixing the lattice constants, cell geometries, and M−X bond distances and X−M−X angles of BX_6 octahedra prior to optimizing the BX_6 tilting magnitudes.[42] These perovskite structure properties are not constrained during the DFT optimization of SPuDS-generated Glazer tilts but are rather allowed to change simultaneously to minimize the energy. There is no guarantee that $R_{0,GS\text{-}DFT}$s, when used by SPuDS's constrained optimization algorithm, will recover DFT structures that are freely optimized without these imposed constraints. Additionally, the SPuDS algorithm over-rotates BX_6 octahedra for the a-a-a- Glazer tilt phase, resulting in unrealistic A−O and O−O bond lengths.[42] This issue cannot be corrected by the cation−anion pair $R_{0,GS\text{-}DFT}$s reported herein, although anion−anion pair $R_{0,GS\text{-}DFT}$ values derived using our framework that restrict unrealistic O−O distances could rectify this issue. Consequently, we cannot recommend $R_{0,GS\text{-}DFT}$s derived herein for the SPuDS algorithm until these SPuDS-specific issues have been addressed.

CONCLUSIONS

Our results demonstrate that cation−anion pair BVM bond length parameters $R_{0,GS\text{-}DFT}$s, parameterized to minimize $GII_{GS\text{-}DFT}$ while enforcing a linear relationship between the GIIs and DFT energies of competing structures, accurately predict the structural and energetic trends of ABO_3 perovskite oxides. Composition-specific $R_{0,GS\text{-}DFT}$ parameters correctly predict the DFT ground states from GIIs for 174 of 188 compositions considered, confirming that bond-valence model GIIs with optimal R_0 parameters reliably capture DFT relative stabilities in ABO_3 perovskite oxides. The 59 general cation−anion $R_{0,GS\text{-}DFT}$ parameters derived using our optimization framework accurately capture the DFT ground states from GIIs for 135 of 188 compositions, or 152 of 188 compositions if the DFT ground state structure is one of the two structures with the lowest GII values. For lanthanide-transition-metal perovskite oxide compositions $(Ln^{3+})(Tm_{3d})O_3$ that are of particular interest to the STCH community, the DFT ground state is correctly predicted for 75 of 89 compositions. Furthermore, BVM ideal bond length $R_{0,GS\text{-}DFT}$ parameters predict A−O and B−O bond distances in better agreement with the DFT ground state structures than $R_{0,rmsd}$s, resulting in t_{bv}s that predict greater energetic stabilization due to octahedral tilting, in agreement with DFT. BVM parameterization by the constrained minimization of $GII_{GS\text{-}DFT}$s therefore markedly improves upon existing parameterization frameworks and could enable a significant reduction in the computational expense of high-throughput DFT investigations of compelling compositional spaces, such as the computational screening of perovskite oxides as STCH redox mediators.

METHODS

Dataset Generation. ABO_3 compositions considered in the present investigation were taken from the 232 experimentally observed ternary perovskite oxide compositions tabulated by Zhang et al.[43] The 15 experimentally observed oxide compositions that include radioactive elements and iodine were excluded from the present investigation. The remaining 217 compositions were assigned oxidation states using the algorithm reported by Bartel et al.[44] and input to SPuDS (DOS version >2.20.08.06) using a custom python wrapper; this python wrapper is publicly available for download at https://github.com/zaba1157/PySPuDS. Of the 217 compositions considered, 29 compositions exhibit a bond-valence Goldschmidt tolerance factor, t_{bv}, greater than 1 when computed with the default SPuDS BVM parameters. For compositions with $t_{bv} > 1.0$, SPuDS only outputs the cubic a0a0a0 phase,[42] rather than the a0a0a0 phase as well as the 10 additional Glazer octahedral tilt systems output for compositions exhibiting $t_{bv} \leq 1.0$. Because the a0a0a0 phase alone cannot represent the GII−DFT energy relationship across many different competing structures, and therefore parameterization principles 2 and 3 cannot be benchmarked for these compositions, these 29 compositions were excluded from the final dataset. This resulted in a final dataset of 188 compositions, with structures that were generated in the cubic phase as well as 10 competing Glazer octahedral tilt systems and optimized using DFT. The dataset is composed of 59 unique cation−anion pairs, the $R_{0,GS\text{-}DFT}$ and $R_{0,rmsd}$ parameters of which were fit using the parameterization schemes described herein.

DFT Parameters and Materials Project Compatibility. GGA + U DFT calculations were performed using the Vienna ab initio simulation program (VASP 5.4.1)[45−47] with the Perdew−Burke−Ernzerhof (PBE)[48] exchange−correlation functional and periodic boundary conditions utilizing projector-augmented wave pseudopotentials.[49,50] All calculations are compatible with the Materials Project (MP) database, which tabulates the structures and energies of >140,000 inorganic materials computed at the PBE (GGA + U) level of theory. Hubbard U parameters for the elements Mn, Fe, Co, Cr, Mo, W, V, and Ni were taken from the pymatgen[51] python package's MPRelaxSet, which tabulates the U parameters calibrated using the approach described by Wang et al.[52] Decomposition enthalpies, ΔH_d^{DFT}, have MP-specific corrections, as described by pymatgen's MP compatibility settings. These include oxide-specific corrections based on the structure type (i.e., oxide, peroxide, superoxide, or ozonide). As discussed in https://docs.materialsproject.org/methodology/total-energies/, the coordination environments of oxygen anions present in these materials determine this classification. The ΔH_d^{DFT} values of DFT-optimized ABO_3 competing structures in our dataset were calculated relative to their lowest enthalpy decomposition products, as tabulated in the MP database.

The electronic wave functions were expanded in a plane wave basis set with an energy cutoff of 520 eV. The Brillouin zones were sampled during geometry optimizations using the Monkhorst−Pack algorithm to automatically generate a Γ-point-centered k-point mesh with a grid density of at least 1000/(atoms/unit cell). The specific pseudopotentials used are consistent with pymatgen's MPRelaxSet, which is the default relaxation parameter set for the MP. Two consecutive spin-polarized relaxations were performed with magnetic ions initialized in a high-spin ferromagnetic configuration where

https://doi.org/10.1021/acs.jctc.1c01113
J. Chem. Theory Comput. 2022, 18, 3257−3267

initial spin magnitudes were species-specific, as dictated by the default MP spin parameters. For geometry optimizations, the electronic self-consistent field steps were converged to within 1×10^{-6} eV, while forces were converged to within 0.01 eV/Å. Single-point magnetic sampling using pymatgen's MagneticStructureEnumerator module was performed on all structures with magnetic elements. This ensures that the DFT energies of structures were reported in their lowest energy spin configurations and that $R_{0,\text{GS-DFT}}$s also captured stabilization due to magnetism. We do not optimize separate $R_{0,\text{GS-DFT}}$ or $R_{0,\text{rmsd}}$ parameters for cations exhibiting different magnetic moments in DFT, for example, high-spin $Co^{3+}-O^{2-}$ versus low-spin $Co^{3+}-O^{2-}$.

GII Calculation. GIIs were computed using the GIICalculator class, which calculates the GIIs of oxidation-state-decorated pymatgen.core.structure.Structure objects and can be found at https://github.com/rymo1354/gii_minimization. GIICalculator exploits the space group symmetries and employs bond-valence parameter caching to rapidly compute GIIs, thereby significantly reducing the runtime of GII calculations and R_0 parameterization relative to the existing python-based GII calculators.[53,54] GIICalculator can compute GIIs using parameters tabulated by Brown[55] or from user-specific parameters. Additionally, nearest neighbors for the GII calculation can be identified using a fixed neighbor radius search or by using pymatgen's CrystalNN, which chooses neighbors based upon their oxidation states using a modified Voronoi algorithm. CrystalNN is used in the present work because it does not require a user-specified nearest neighbor radius, which can be composition- and system (i.e., oxide, halide, etc.)-specific.

Optimization Algorithm. Parameterizations of general and composition-specific $R_{0,\text{GS-DFT}}$ and $R_{0,\text{rmsd}}$ parameters were performed using the trust-region-constrained algorithm,[56] as implemented in the open-source SciPy package.[57] This algorithm minimizes a function with the following form

$$\min f(x)$$

such that $c_l \leq c(x) \leq c_u$, $x_l \leq x \leq x_u$, where $c(x)$ is a nonlinear inequality constraint, and $x_l \leq x \leq x_u$ specifies any bounds imposed on the vector x to be optimized. Simultaneous parameterization of R_0s and Bs can be performed using the trust-region-constrained algorithm, as both parameters can be expressed by the optimization parameter vector x. This algorithm performs a local minimization, which makes it sensitive to the initial parameters passed. It can also identify optimized parameters that violate imposed constraints if the user-specified trust region radius tolerance is exceeded, which must be considered during parameterization. Otherwise, convergence is reached if the Lagrangian gradient tolerance or barrier tolerance conditions are achieved. For the optimizations performed in this work, the Lagrangian gradient tolerance was fixed to 0.001, whereas the trust radius and barrier tolerance were fixed to 0.01. Python implementation of the trust-region-constrained algorithm used herein for BVM parameterization is available at https://github.com/rymo1354/gii_minimization.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jctc.1c01113.

Composition-specific $R_{0,\text{GS-DFT}}$ parameters; initial estimates of $R_{0,\text{GS-DFT}}$s for parameterization derived from Shannon radii; evaluation of the Pearson coefficient constraint constant, C; discussion regarding the correct ordering of energetics by GIIs; composition-specific $R_{0,\text{GS-DFT}}$ parameter distributions and discussion explaining why some distributions are broader than others; example of weighted general $R_{0,\text{GS-DFT}}$s parameterization; $R_{0,\text{GS-DFT}}$s re-parameterized for a subset of 11 Glazer tilts; and comparison of general $R_{0,\text{GS-DFT}}$s to $R_{0,\text{rmsd}}$s parameters derived from our dataset (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Charles B. Musgrave − *Department of Chemical and Biological Engineering, University of Colorado Boulder, Boulder, Colorado 80309, United States; Materials Science and Engineering Program, University of Colorado Boulder, Boulder, Colorado 80309, United States;* orcid.org/0000-0002-5732-3180; Email: charles.musgrave@colorado.edu

Authors

Ryan J. Morelock − *Department of Chemical and Biological Engineering, University of Colorado Boulder, Boulder, Colorado 80309, United States;* orcid.org/0000-0001-8368-8258

Zachary J. L. Bare − *Department of Chemical and Biological Engineering, University of Colorado Boulder, Boulder, Colorado 80309, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jctc.1c01113

Author Contributions

R.J.M. and Z.J.L.B. contributed equally. R.J.M. conducted the parameter optimization and drafted the manuscript. Z.J.L.B. generated the dataset. C.B.M supervised the DFT calculations and optimization and assisted with the editing and writing of the manuscript.

Funding

This work was supported by the U.S. Department of Energy (DOE), Office of Energy Efficiency and Renewable Energy (EERE), Hydrogen and Fuel Cell Technologies Office (HFTO), and specifically the HydroGEN Advanced Water Splitting Materials Consortium, established as part of the Energy Materials Network under this same office (award DE-EE0008088). C.B.M., Z.J.L.B., and R.J.M. also acknowledge the support from the National Science Foundation (awards NSF CHEM-1800592 and CBET-2016225). R.J.M. acknowledges the support from a University of Colorado Boulder's Materials for Energy Conversion and Sustainability Graduate Assistance in Areas of National Need (GAANN) fellowship, Department of Education (award no. P200A180012). The views expressed in this article do not necessarily represent the views of the U.S. Department of Energy or the United States Government.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors would like to acknowledge Dr. Christopher Sutton for his helpful discussions.

■ ABBREVIATIONS

BVM, bond-valence method; GII, global instability index; DFT, density functional theory; rmsd, root-mean-square deviation; STCH, solar thermochemical hydrogen; MP, Materials Project; GS, ground state; SPuDS, Structure Prediction and Diagnostic Software

■ REFERENCES

(1) Pauling, L. The Principles Determining the Structure of Complex Ionic Crystals. *J. Am. Chem. Soc.* **1929**, *51*, 1010−1026.

(2) Brown, I. D. 14-The Bond-Valence Method: An Empirical Approach to Chemical Structure and Bonding. In *Industrial Chemistry Library*; O'Keeffe, M., Navrotsky, A., Eds.; Elsevier, 1981; Vol. 2, pp 1−30.

(3) Pauling, L. Atomic Radii and Interatomic Distances in Metals. *J. Am. Chem. Soc.* **1947**, *69*, 542−553.

(4) Donnay, G.; Allmann, R. How to Recognize O^{2-}, OH^-, and H_2O in Crystal Structures Determined by x-Rays. *Am. Mineral.* **1970**, *55*, 1003.

(5) Zachariasen, W. H. Bond Lengths in Oxygen and Halogen Compounds of d and f Elements. *J. Less-Common Met.* **1978**, *62*, 1−7.

(6) Brown, I. D. Recent Developments in the Methods and Applications of the Bond Valence Model. *Chem. Rev.* **2009**, *109*, 6858−6919.

(7) Etxebarria, I.; Perez-Mato, J. M.; García, A.; Blaha, P.; Schwarz, K.; Rodriguez-Carvajal, J. Comparison of Empirical Bond-Valence and First-Principles Energy Calculations for a Complex Structural Instability. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2005**, *72*, 174108.

(8) Yamada, I.; Takamatsu, A.; Ikeno, H. Complementary Evaluation of Structure Stability of Perovskite Oxides Using Bond-Valence and Density-Functional-Theory Calculations. *Sci. Technol. Adv. Mater.* **2018**, *19*, 101−107.

(9) Wood, R. M.; Palenik, G. J. Bond Valence Sums in Coordination Chemistry. A Simple Method for Calculating the Oxidation State of Cobalt in Complexes Containing Only Co-O Bonds. *Inorg. Chem.* **1998**, *37*, 4149−4151.

(10) Brese, N. E.; O'Keeffe, M. Bond-Valence Parameters for Solids. *Acta Crystallogr., Sect. B: Struct. Sci.* **1991**, *47*, 192−197.

(11) Muhich, C. L.; Evanko, B. W.; Weston, K. C.; Lichty, P.; Liang, X.; Martinek, J.; Musgrave, C. B.; Weimer, A. W. Efficient Generation of H_2 by Splitting Water with an Isothermal Redox Cycle. *Science* **2013**, *341*, 540−542.

(12) McDaniel, A. H.; Ambrosini, A.; Coker, E. N.; Miller, J. E.; Chueh, W. C.; O'Hayre, R.; Tong, J. Nonstoichiometric Perovskite Oxides for Solar Thermochemical H_2 and CO Production. *Energy Proc.* **2014**, *49*, 2009.

(13) Rao, C. N. R.; Dey, S. Solar Thermochemical Splitting of Water to Generate Hydrogen. *Proc. Natl. Acad. Sci. U.S.A.* **2017**, *114*, 13385−13393.

(14) Kubicek, M.; Bork, A. H.; Rupp, J. L. M. Perovskite Oxides − a Review on a Versatile Material Class for Solar-to-Fuel Conversion Processes. *J. Mater. Chem. A* **2017**, *5*, 11983−12000.

(15) Castelli, I. E.; Olsen, T.; Datta, S.; Landis, D. D.; Dahl, S.; Thygesen, K. S.; Jacobsen, K. W. Computational Screening of Perovskite Metal Oxides for Optimal Solar Light Capture. *Energy Environ. Sci.* **2012**, *5*, 5814−5819.

(16) Emery, A. A.; Wolverton, C. High-Throughput DFT Calculations of Formation Energy, Stability and Oxygen Vacancy Formation Energy of ABO_3 Perovskites. *Sci. Data* **2017**, *4*, 170153.

(17) Sai Gautam, G.; Stechel, E. B.; Carter, E. A. Exploring Ca−Ce−M−O (M = 3d Transition Metal) Oxide Perovskites for Solar Thermochemical Applications. *Chem. Mater.* **2020**, *32*, 9964−9982.

(18) Bare, Z. J. L.; Morelock, R. J.; Musgrave, C. B. A Computational Framework to Accelerate the Discovery of Perovskites for Solar Thermochemical Hydrogen Production; Identification of Gd Perovskite Oxide Redox Mediators. *Adv. Funct. Mater.* **2022**, 2200201.

(19) Glazer, A. M. The Classification of Tilted Octahedra in Perovskites. *Acta Crystallogr., Sect. B: Struct. Crystallogr. Cryst. Chem.* **1972**, *28*, 3384−3392.

(20) Zhao, T.; Daniels, L. M.; Slater, B.; Rosseinsky, M. J.; Corà, F. Effects of Octahedral Tilting on Band Structure and Thermoelectric Power Factor of Titanate Perovskites: A First-Principles Study on $SrTiO_3$. *J. Phys. Chem. C* **2020**, *124*, 13045−13052.

(21) Xie, N.; Zhang, J.; Raza, S.; Zhang, N.; Chen, X.; Wang, D. Generation of Low-Symmetry Perovskite Structures for Ab Initio Computation. *J. Phys.: Condens. Matter* **2020**, *32*, 315901.

(22) Bartel, C. J.; Clary, J. M.; Sutton, C.; Vigil-Fowler, D.; Goldsmith, B. R.; Holder, A. M.; Musgrave, C. B. Inorganic Halide Double Perovskites with Optoelectronic Properties Modulated by Sublattice Mixing. *J. Am. Chem. Soc.* **2020**, *142*, 5135−5145.

(23) Brown, I. D. What Is the Best Way to Determine Bond-Valence Parameters? *IUCrJ* **2017**, *4*, 514−515.

(24) Gagné, O. C.; Hawthorne, F. C. Comprehensive Derivation of Bond-Valence Parameters for Ion Pairs Involving Oxygen. *Acta Crystallogr. B Struct. Sci. Cryst. Eng. Mater.* **2015**, *71*, 562−578.

(25) Lufaso, M. W.; Woodward, P. M. Prediction of the Crystal Structures of Perovskites Using the Software Program SPuDS. *Acta Crystallogr., Sect. B: Struct. Sci.* **2001**, *57*, 725−738.

(26) Jain, A.; Ong, S. P.; Hautier, G.; Chen, W.; Richards, W. D.; Dacek, S.; Cholia, S.; Gunter, D.; Skinner, D.; Ceder, G.; Persson, K. A. Commentary: The Materials Project: A Materials Genome Approach to Accelerating Materials Innovation. *APL Mater.* **2013**, *1*, 011002.

(27) Brown, I. D.; Altermatt, D. Bond-Valence Parameters Obtained from a Systematic Analysis of the Inorganic Crystal Structure Database. *Acta Crystallogr., Sect. B: Struct. Sci.* **1985**, *41*, 244−247.

(28) Shannon, R. D. Revised Effective Ionic Radii and Systematic Studies of Interatomic Distances in Halides and Chalcogenides. *Acta Crystallogr., Sect. A: Cryst. Phys., Diffr., Theor. Gen. Crystallogr.* **1976**, *32*, 751−767.

(29) Freedman, D. A. *Statistical Models: Theory and Practice*; Cambridge University Press, 2009.

(30) Hautier, G.; Ong, S. P.; Jain, A.; Moore, C. J.; Ceder, G. Accuracy of Density Functional Theory in Predicting Formation Energies of Ternary Oxides from Binary Oxides and Its Implication on Phase Stability. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2012**, *85*, 155208.

(31) Park, J. E.; Bare, Z. J. L.; Morelock, R. J.; Rodriguez, M. A.; Ambrosini, A.; Musgrave, C. B.; McDaniel, A. H.; Coker, E. N. Computationally Accelerated Discovery and Experimental Demonstration of $Gd_{0.5}La_{0.5}Co_{0.5}Fe_{0.5}O_3$ for Solar Thermochemical Hydrogen Production. *Front. Energy Res.* **2021**, *9*, 750600.

(32) Morelock, R. J.; Tran, J.; Bare, Z. J. L.; Trindell, J. A.; McDaniel, A. H.; Weimer, A. W.; Musgrave, C. B. Computationally Guided Discovery of Mixed Mn/Ni Perovskites for Solar Thermochemical Hydrogen Production at High H_2 Conversion. In progress, **2022**.

(33) Orlov, Y. S.; Solovyov, L. A.; Dudnikov, V. A.; Fedorov, A. S.; Kuzubov, A. A.; Kazak, N. V.; Voronov, V. N.; Vereshchagin, S. N.; Shishkina, N. N.; Perov, N. S.; Lamonova, K. V.; Babkin, R. Y.; Pashkevich, Y. G.; Anshits, A. G.; Ovchinnikov, S. G. Structural Properties and High-Temperature Spin and Electronic Transitions in $GdCoO_3$: Experiment and Theory. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2013**, *88*, 235105.

(34) Goldschmidt, V. M. Die Gesetze Der Krystallochemie. *Naturwissenschaften* **1926**, *14*, 477−485.

(35) Urusov, V. S. Problem of Optimization of Bond Valence Model Parameters (as Exemplified by Managanese in Different Oxidation States). *Dokl. Phys. Chem.* **2006**, *408*, 152−155.

(36) Adams, S. Relationship between Bond Valence and Bond Softness of Alkali Halides and Chalcogenides. *Acta Crystallogr., Sect. B: Struct. Sci.* **2001**, *57*, 278−287.

(37) Wander, M. C. F.; Bickmore, B. R.; Davis, M.; Johansen, W. J.; Andros, C.; Lind, L. The Use of Cation-Cation and Anion-Anion Bonds to Augment the Bond-Valence Model. *Am. Mineral.* **2015**, *100*, 148−159.

(38) Harris, S. E.; Orpen, A. G.; Bruno, I. J.; Taylor, R. Factors Affecting D-Block Metal-Ligand Bond Lengths: Toward an Automated Library of Molecular Geometry for Metal Complexes. *J. Chem. Inf. Model.* **2005**, *45*, 1727−1748.

https://doi.org/10.1021/acs.jctc.1c01113
J. Chem. Theory Comput. 2022, 18, 3257−3267

(39) Cai, Y.; Xie, W.; Teng, Y. T.; Harikesh, P. C.; Ghosh, B.; Huck, P.; Persson, K. A.; Mathews, N.; Mhaisalkar, S. G.; Sherburne, M.; Asta, M. High-Throughput Computational Study of Halide Double Perovskite Inorganic Compounds. *Chem. Mater.* **2019**, *31*, 5392−5401.
(40) Kar, M.; Körzdörfer, T. Computational High Throughput Screening of Inorganic Cation Based Halide Perovskites for Perovskite Only Tandem Solar Cells. *Mater. Res. Express* **2020**, *7*, 055502.
(41) Xie, S. R.; Kotlarz, P.; Hennig, R. G.; Nino, J. C. Machine Learning of Octahedral Tilting in Oxide Perovskites by Symbolic Classification with Compressed Sensing. *Comput. Mater. Sci.* **2020**, *180*, 109690.
(42) Lufaso, M. W.; Woodward, P. *SPuDS Users Guide*, 2001.
(43) Zhang, H.; Li, N.; Li, K.; Xue, D. Structural Stability and Formability of ABO_3-Type Perovskite Compounds. *Acta Crystallogr., Sect. B: Struct. Sci.* **2007**, *63*, 812−818.
(44) Bartel, C. J.; Sutton, C.; Goldsmith, B. R.; Ouyang, R.; Musgrave, C. B.; Ghiringhelli, L. M.; Scheffler, M. New Tolerance Factor to Predict the Stability of Perovskite Oxides and Halides. *Sci. Adv.* **2019**, *5*, No. eaav0693.
(45) Kresse, G.; Hafner, J. Norm-Conserving and Ultrasoft Pseudopotentials for First-Row and Transition Elements. *J. Phys.: Condens. Matter* **1994**, *6*, 8245.
(46) Kresse, G.; Furthmüller, J. Efficiency of Ab-Initio Total Energy Calculations for Metals and Semiconductors Using a Plane-Wave Basis Set. *Comput. Mater. Sci.* **1996**, *6*, 15−50.
(47) Kresse, G.; Furthmüller, J. Efficient Iterative Schemes for Ab Initio Total-Energy Calculations Using a Plane-Wave Basis Set. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1996**, *54*, 11169−11186.
(48) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865−3868.
(49) Kresse, G.; Hafner, J. Ab Initio Molecular Dynamics for Liquid Metals. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1993**, *47*, 558−561.
(50) Kresse, G.; Joubert, D. From Ultrasoft Pseudopotentials to the Projector Augmented-Wave Method. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1999**, *59*, 1758−1775.
(51) Ong, S. P.; Richards, W. D.; Jain, A.; Hautier, G.; Kocher, M.; Cholia, S.; Gunter, D.; Chevrier, V. L.; Persson, K. A.; Ceder, G. Python Materials Genomics (Pymatgen): A Robust, Open-Source Python Library for Materials Analysis. *Comput. Mater. Sci.* **2013**, *68*, 314−319.
(52) Wang, L.; Maxisch, T.; Ceder, G. Oxidation Energies of Transition Metal Oxides within the GGA+U Framework. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2006**, *73*, 195107.
(53) Ward, L.; Dunn, A.; Faghaninia, A.; Zimmermann, N. E. R.; Bajaj, S.; Wang, Q.; Montoya, J.; Chen, J.; Bystrom, K.; Dylla, M.; Chard, K.; Asta, M.; Persson, K. A.; Snyder, G. J.; Foster, I.; Jain, A. Matminer: An Open Source Toolkit for Materials Data Mining. *Comput. Mater. Sci.* **2018**, *152*, 60−69.
(54) Salinas-Sanchez, A.; Garcia-Muñoz, J. L.; Rodriguez-Carvajal, J.; Saez-Puche, R.; Martinez, J. L. Structural Characterization of R_2BaCuO_5 (R = Y, Lu, Yb, Tm, Er, Ho, Dy, Gd, Eu and Sm) Oxides by X-Ray and Neutron Diffraction. *J. Solid State Chem.* **1992**, *100*, 201−211.
(55) Bond valence parameters https://www.iucr.org/resources/data/datasets/bond-valence-parameters (accessed Dec 10, 2021).
(56) Byrd, R. H.; Hribar, M. E.; Nocedal, J. An Interior Point Algorithm for Large Scale Nonlinear Programming. *SIAM J. Optim.* **1997**, *9*, 877.
(57) Virtanen, P.; Gommers, R.; Oliphant, T. E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J.; van der Walt, S. J.; Brett, M.; Wilson, J.; Millman, K. J.; Mayorov, N.; Nelson, A. R. J.; Jones, E.; Kern, R.; Larson, E.; Carey, C. J.; Polat, İ.; Feng, Y.; Moore, E. W.; VanderPlas, J.; Laxalde, D.; Perktold, J.; Cimrman, R.; Henriksen, I.; Quintero, E. A.; Harris, C. R.; Archibald, A. M.; Ribeiro, A. H.; Pedregosa, F.; van Mulbregt, P.; Vijaykumar, A.; Bardelli, A. P.; Rothberg, A.; Hilboll, A.; Kloeckner, A.; Scopatz, A.; Lee, A.; Rokem, A.; Woods, C. N.; Fulton, C.; Masson, C.; Häggström, C.; Fitzgerald, C.; Nicholson, D. A.; Hagen, D. R.; Pasechnik, D. V.; Olivetti, E.; Martin, E.; Wieser, E.; Silva, F.; Lenders, F.; Wilhelm, F.; Young, G.; Price, G. A.; Ingold, G.-L.; Allen, G. E.; Lee, G. R.; Audren, H.; Probst, I.; Dietrich, J. P.; Silterra, J.; Webber, J. T.; Slavič, J.; Nothman, J.; Buchner, J.; Kulick, J.; Schönberger, J. L.; de Miranda Cardoso, J. V.; Reimer, J.; Harrington, J.; Rodríguez, J. L. C.; Nunez-Iglesias, J.; Kuczynski, J.; Tritz, K.; Thoma, M.; Newville, M.; Kümmerer, M.; Bolingbroke, M.; Tartre, M.; Pak, M.; Smith, N. J.; Nowaczyk, N.; Shebanov, N.; Pavlyk, O.; Brodtkorb, P. A.; Lee, P.; McGibbon, R. T.; Feldbauer, R.; Lewis, S.; Tygier, S.; Sievert, S.; Vigna, S.; Peterson, S.; More, S.; Pudlik, T.; Oshima, T.; Pingel, T. J.; Robitaille, T. P.; Spura, T.; Jones, T. R.; Cera, T.; Leslie, T.; Zito, T.; Krauss, T.; Upadhyay, U.; Halchenko, Y. O.; Vázquez-Baeza, Y. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nat. Methods* **2020**, *17*, 261−272.

https://doi.org/10.1021/acs.jctc.1c01113
J. Chem. Theory Comput. 2022, 18, 3257−3267